

MEDLINE [76150089](#)
PUBMED [768781](#)
REFERENCE 8 (bases 1242 to 1268)
AUTHORS Heyneker, H.L., Shine, J., Goodman, H.M., Boyer, H.W., Rosenberg, J.,
Dickerson, R.E., Narang, S.A., Itakura, K., Lin, S. and Riggs, A.D.
TITLE Synthetic lac operator DNA is functional in vivo
JOURNAL Nature 263 (5580), 748-752 (1976)
MEDLINE [77056376](#)
PUBMED [1069185](#)
REFERENCE 9 (sites)
AUTHORS Dickson, R.C., Abelson, J. and Johnson, P.
TITLE Nucleotide sequence changes produced by mutations in the lac
promoter of Escherichia coli
JOURNAL J. Mol. Biol. 111 (1), 65-75 (1977)
MEDLINE [77168230](#)
PUBMED [323498](#)
REFERENCE 10 (bases 51 to 264)
AUTHORS Steege, D.A.
TITLE 5'-Terminal nucleotide sequence of Escherichia coli lactose
repressor mRNA: features of translational initiation and
reinitiation sites
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 74 (10), 4163-4167 (1977)
MEDLINE [78052881](#)
PUBMED [337294](#)
REFERENCE 11 (bases 1 to 81)
AUTHORS Calos, M.P.
TITLE DNA sequence for a low-level promoter of the lac repressor gene and
an 'up' promoter mutation
JOURNAL Nature 274 (5673), 762-765 (1978)
MEDLINE [78246990](#)
PUBMED [355890](#)
REFERENCE 12 (bases 49 to 1161)
AUTHORS Farabaugh, P.J.
TITLE Sequence of the lacI gene
JOURNAL Nature 274 (5673), 765-769 (1978)
MEDLINE [78246991](#)
PUBMED [355891](#)
REFERENCE 13 (sites)
AUTHORS Miller, J.H., Coulondre, C. and Farabaugh, P.J.
TITLE Correlation of nonsense sites in the lacI gene with specific codons
in the nucleotide sequence
JOURNAL Nature 274 (5673), 770-775 (1978)
MEDLINE [78246992](#)
PUBMED [355892](#)
REFERENCE 14 (sites)
AUTHORS Calos, M.P. and Miller, J.H.
TITLE DNA sequence alteration resulting from a mutation impairing
promoter function in the lac repressor gene
JOURNAL Mol. Gen. Genet. 178 (1), 225-227 (1980)
MEDLINE [80209248](#)
PUBMED [6770231](#)
REFERENCE 15 (bases 4306 to 5804)
AUTHORS Buchel, D.E., Gronenborn, B. and Muller-Hill, B.
TITLE Sequence of the lactose permease gene
JOURNAL Nature 283 (5747), 541-545 (1980)
MEDLINE [80120651](#)
PUBMED [6444453](#)
REFERENCE 16 (sites)
AUTHORS Miller, J.H., Calos, M.P. and Galas, D.J.
TITLE Genetic and sequencing studies of the specificity of transposition
into the lac region of E. coli
JOURNAL Cold Spring Harb. Symp. Quant. Biol. 45 Pt 1, 243-257 (1981)
MEDLINE [82049502](#)
PUBMED [6271472](#)

- REFERENCE 17 (sites)
AUTHORS Chenchick,A., Beabealashvilli,R. and Mirzabekov,A.
TITLE Topography of interaction of Escherichia coli RNA polymerase subunits with lac UV5 promoter
JOURNAL FEBS Lett. 128 (1), 46-50 (1981)
MEDLINE [82004657](#)
PUBMED [7023981](#)
- REFERENCE 18 (sites)
AUTHORS Betz,J.L. and Sadler,J.R.
TITLE Variants of a cloned synthetic lactose operator. I. A palindromic dimer lactose operator derived from one stand of the cloned 40-base pair operator
JOURNAL Gene 13 (1), 1-12 (1981)
MEDLINE [81213459](#)
PUBMED [7016667](#)
- REFERENCE 19 (sites)
AUTHORS Sadler,J.R. and Tecklenburg,M.
TITLE Cloning and characterization of the natural lactose operator
JOURNAL Gene 13 (1), 13-23 (1981)
MEDLINE [81213463](#)
PUBMED [6263752](#)
- REFERENCE 20 (sites)
AUTHORS Betz,J.L. and Sadler,J.R.
TITLE Variants of a cloned synthetic lactose operator. II. Chloramphenicol-resistant revertants retaining a lactose operator in the CAT gene of plasmid pBR325
JOURNAL Gene 15 (2-3), 187-200 (1981)
MEDLINE [82051311](#)
PUBMED [6271642](#)
- REFERENCE 21 (sites)
AUTHORS Calos,M.P. and Miller,J.H.
TITLE The DNA sequence change resulting from the IQ1 mutation, which greatly increases promoter strength
JOURNAL Mol. Gen. Genet. 183 (3), 559-560 (1981)
MEDLINE [82147746](#)
PUBMED [7038381](#)
- REFERENCE 22 (sites)
AUTHORS Mieschendahl,M., Buchel,D., Bocklage,H. and Muller-Hill,B.
TITLE Mutations in the lacY gene of Escherichia coli define functional organization of lactose permease
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 78 (12), 7652-7656 (1981)
MEDLINE [82150928](#)
PUBMED [6278484](#)
- REFERENCE 23 (sites)
AUTHORS Russell,D.R. and Bennett,G.N.
TITLE Construction and analysis of in vivo activity of E. coli promoter hybrids and promoter mutants that alter the -35 to -10 spacing
JOURNAL Gene 20 (2), 231-243 (1982)
MEDLINE [83158761](#)
PUBMED [6299890](#)
- REFERENCE 24 (sites)
AUTHORS Horowitz,H. and Platt,T.
TITLE A termination site for LacI transcription is between the CAP site and the lac promoter
JOURNAL J. Biol. Chem. 257 (19), 11740-11746 (1982)
MEDLINE [83007251](#)
PUBMED [6288696](#)
- REFERENCE 25 (sites)
AUTHORS Klein,R.D. and Wells,R.D.
TITLE Effects of neighboring DNA homopolymers on the biochemical and physical properties of the Escherichia coli lactose promoter. I. Cloning and characterization studies
JOURNAL J. Biol. Chem. 257 (21), 12954-12961 (1982)
MEDLINE [83030833](#)

PUBMED [6290487](#)
REFERENCE 26 (bases 1183 to 1291)
AUTHORS Weiher, H. and Schaller, H.
TITLE Segment-specific mutagenesis: extensive mutagenesis of a lac promoter/operator element
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 79 (5), 1408-1412 (1982)
MEDLINE [82174608](#)
PUBMED [7041119](#)
REFERENCE 27 (sites)
AUTHORS Van Dyke, M.W. and Dervan, P.B.
TITLE Footprinting with MPE.Fe(II). Complementary-strand analyses of distamycin- and actinomycin-binding sites on heterogeneous DNA
JOURNAL Cold Spring Harb. Symp. Quant. Biol. 47 Pt 1, 347-353 (1983)
MEDLINE [83233528](#)
PUBMED [6305557](#)
REFERENCE 28 (bases 1287 to 4364)
AUTHORS Kalnins, A., Otto, K., Ruther, U. and Muller-Hill, B.
TITLE Sequence of the lacZ gene of Escherichia coli
JOURNAL EMBO J. 2 (4), 593-597 (1983)
MEDLINE [84028567](#)
PUBMED [6313347](#)
REFERENCE 29 (sites)
AUTHORS Cone, K.C., Sellitti, M.A. and Steege, D.A.
TITLE Lac repressor mRNA transcription terminates in vivo in the lac control region
JOURNAL J. Biol. Chem. 258 (18), 11296-11304 (1983)
MEDLINE [83291093](#)
PUBMED [6309841](#)
REFERENCE 30 (sites)
AUTHORS Sadler, J.R., Sasmor, H. and Betz, J.L.
TITLE A perfectly symmetric lac operator binds the lac repressor very tightly
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 80 (22), 6785-6789 (1983)
MEDLINE [84070714](#)
PUBMED [6316325](#)
REFERENCE 31 (sites)
AUTHORS Glickman, B.W. and Ripley, L.S.
TITLE Structural intermediates of deletion mutagenesis: a role for palindromic DNA
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 81 (2), 512-516 (1984)
MEDLINE [84119517](#)
PUBMED [6582506](#)
REFERENCE 32 (sites)
AUTHORS Spassky, A., Kirkegaard, K. and Buc, H.
TITLE Changes in the DNA structure of the lac UV5 promoter during formation of an open complex with Escherichia coli RNA polymerase
JOURNAL Biochemistry 24 (11), 2723-2731 (1985)
MEDLINE [85280412](#)
PUBMED [3896305](#)
REFERENCE 33 (sites)
AUTHORS Straney, D.C. and Crothers, D.M.
TITLE Intermediates in transcription initiation from the E. coli lac UV5 promoter
JOURNAL Cell 43 (2 Pt 1), 449-459 (1985)
MEDLINE [86079527](#)
PUBMED [2416465](#)
REFERENCE 34 (sites)
AUTHORS Looman, A.C., de Gruyter, M., Vogelaar, A. and van Knippenberg, P.H.
TITLE Effects of heterologous ribosomal binding sites on the transcription and translation of the lacZ gene of Escherichia coli
JOURNAL Gene 37 (1-3), 145-154 (1985)
MEDLINE [86031346](#)
PUBMED [3932130](#)
REFERENCE 35 (sites)

AUTHORS Mandecki,W., Goldman,R.A., Powell,B.S. and Caruthers,M.H.
TITLE lac Up-promoter mutants with increased homology to the consensus promoter sequence
JOURNAL J. Bacteriol. 164 (3), 1353-1355 (1985)
MEDLINE 86059235
PUBMED 2999082
REFERENCE 36 (sites)
AUTHORS Malamy,M.H., Rahaim,P.T., Hoffman,C.S., Baghdoyan,D., O'Connor,M.B. and Miller,J.F.
TITLE A frameshift mutation at the junction of an IS1 insertion within lacZ restores beta-galactosidase activity via formation of an active lacZ-IS1 fusion protein
JOURNAL J. Mol. Biol. 181 (4), 551-555 (1985)
MEDLINE 85210885
PUBMED 2987506
REFERENCE 37 (bases 5646 to 7477)
AUTHORS Hediger,M.A., Johnson,D.F., Nierlich,D.P. and Zabin,I.
TITLE DNA sequence of the lactose operon: the lacA gene and the transcriptional termination region
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 82 (19), 6414-6418 (1985)
MEDLINE 86016712
PUBMED 3901000
COMMENT Original source text: Escherichia coli DNA; mRNA; clone lambda-h80dlac DNA; clone puk217; pgm8 (see comment).
[3] sites; UV5 mRNA transcripts and operator mutants. [(in) Sund,H. and Blauer,G. (eds.);Protein-Ligand Interactions: 193-207;Walter de] sites; operator mutational analysis. [7] sites; S1 and mung bean nuclease action on operator DNA. [9] sites; class I, II and III promoter mutant analysis. [13] sites; lacI mutant analysis.
[16] sites; Tn5, Tn9 and Tn10 insertion sites in lac region. [14] sites; lacI promoter mutation UJ177.
[18] sites; palindromic dimer operator;. [19] sites; natural operator sequence.
[20] sites; operator mutational analysis. [21] sites; lacI-Q deletion.
[17] sites; RNA polymerase UV5 promoter interaction. [22] sites; lacY mutational analysis.
[24] sites; lacI-promoted transcription termination. [25] sites; wt and UV5 promoter sequence studies. [23] sites; UV5 promoter mutational analysis.
[30] sites; perfectly symmetric operator sequence. [29] sites; lacI mRNA termination site.
[27] sites; distamycin and actinomycin binding to promoter. [31] sites; lacI deletion studies.
[35] sites; promoter mutational studies.
[33] sites; DNAase I studies with promoter sequence. [34] sites; ribosomal binding and translation initiation for lacZ. [36] sites; insertion sequence IS1 integration in lacZ;. [32] sites; DNAase I studies with promoter.
[1] first reports a 27 bp operator(sites 1240-1266) with two-fold symmetries; the operator has also been defined to be bases 1246-1266 or bases 1239-1273 [8]. [(in) Kjeldgaard,N.C. and Maaloe, O.(eds);Control of ribosome synthesis: 138-143;A] explores the ability of lac repressor protein to affect methylation of operator DNA. [8] argues that DNA on both sides of the 21 bp operator (bases 1246-1266) affects repressor binding but that the sequences of this DNA are probably not critical. [5] gives a larger sequence known as the promoter-operator region for the wild-type, whereas [2] and [26] give portions of this region for the mutant strain UV5. Within the promoter region, bases 1162-1199 are identified as the catabolite gene activator protein binding site (cap) and bases 1200-1245 are the RNA polymerase interaction site. [10] reports a

sequence for the 5' end of the lacI (repressor) gene and discusses restart in mutant strains. [11] presents a sequence for the lacI promoter region and identifies an I-Q mutation which enhances lacI transcription approximately ten-fold. [12] gives a complete sequence for lacI which agrees with the known lac repressor sequence. [26] examines the promoter-operator region in the UV5 strain (lacI09) and studies 23 mutant derivatives of this sequence. This sequence agrees with known protein sequences for the lacZ, lacY and lacA enzymes. [15] notes that the fMet codon is not present for lacA and suggests that the 'ttg' codon (5727-5729), which immediately precedes the mature N-terminal asparagine codon, is the start codon. The cds for lacZ, lacY and lacA are included on a single mRNA transcript.

Complete source information:

Escherichia coli DNA [1], [(in) Kjeldgaard, N.C. and Maaloe, O. (eds); Control of ribosome synthesis: 138-143; A], [8], [12], [26]; mRNA [2], [5], [10]; clone
lambda-h80dlac DNA [11], [15]; clone puk217 [28]; pgm8 [37].

FEATURES	Location/Qualifiers
source	1..7477 /organism="Escherichia coli" /mol_type="genomic DNA" /db_xref="taxon:562"
<u>variation</u>	16 /note="c in wild-type; t in 'up' promoter mutant I-Q [11]"
<u>mRNA</u>	51..1230 /note="lacI (repressor) mRNA; preferred in vivo 3' end [12], [29]"
<u>gene</u>	79..1161 /gene="lacI"
<u>CDS</u>	79..1161 /gene="lacI" /note="lac repressor protein (gtg start codon)" /codon_start=1 /transl_table=11 /protein_id="AAA24052.1" /db_xref="GI:146576" /translation="MKPVTLYDVAEYAGVSYQTVSRVVNQASHVSAKTREKVEAAMAE LNYIPNRVAQQLAGKQSLIGVATSSSLALHAPSQIVAAIKSRADQLGASVVVSMVERS GVEACKAAVHNLLAQRVSGLIINYPLDDQDAIAVEAACTNVPALFLDVSDQTPINSII FSHEDGTRLGVEHLVALGHQQIALLAGPLSSVSARLRLAGWHKYLTRNQIQPIAEREG DWSAMSGFQQTMLNEGIVPTAMLVANDQMALGAMRAITESGLRVGADISVVGYYDDT EDSSCYIPPSTTIKQDFRLLGQTSVDRLLQLSQGQAVKGNQLLPVSLVKRKTTLAPNT QTASPRALADSLMQLARQVSRLESGQ"
<u>misc_signal</u>	1162..1199 /note="cap protein binding site"
<u>variation</u>	1183..1186 /note="ttag in wild-type; aatt in strain UV5 [26]"
<u>variation</u>	1209..1211 /note="gct in wild-type; gt in mutant 1305 [5]"
<u>variation</u>	1212 /note="t in wild-type; a in mutant 1241 [5]"
<u>variation</u>	1230 /note="c in wild-type; a in mutant p-r-1a [5]"
<u>variation</u>	1237..1238 /note="gt in wild-type; aa in strain UV5 [26]"
<u>variation</u>	1242..1245 /note="gtgg in wild-type; ttca in synthetic operator [8]"
<u>mRNA</u>	1246..>4358 /note="lacZ mRNA [2], [5]"
<u>misc_signal</u>	1246..1266 /note="lac repressor protein binding site"
<u>variation</u>	1267..1268

variation /note="tc in wild-type; tg in synthetic operator [8]"
1282..1291
/note="ctatgaccat in wild-type; gatccggcca in strain UV5
[26]"
gene 1284..4358
/gene="lacZ"
CDS 1284..4358
/gene="lacZ"
/note="beta-d-galactosidase"
/codon_start=1
/transl_table=11
/protein_id="AAA24053.1"
/db_xref="GI:146577"
/translation="MTMITDSLAVVLQRRDWENPGVTQLNRLAAHPPFASWRNSEEAR
TDRPSQQRLRSINGEWRFAWFPAPEAVPESWLECDLPEADTVVPSNWQMHGYDAPIYT
NVTYPITVNPFFVPTENPTGCYSLTFNVDESWLQEGQTRIIFDGVNSAFHLWCNGRWV
GYGQDSRLPSEFDLSAFLRAGENRLAVMVLRWSDGSYLEDDQDMWRMSGIFRDVSLLLHK
PTTQISDFHVATRFDNDFSRAVLEAEVQMCCELRDYLRVTVSLWQGETQVASGTAPFG
GEIIDERGGYADRVTLRLNVENPKLWSAEIPNLYRAVVELHTADGTLIEAEACDVGFR
EVRIENGLLLLNGKPLLIRGVNRHEHHPLHGQVMDEQTMVQDILLMKQNNFNAVRCSE
YPNHPLWYTLCDRYGLYVVDEANIETHGMVPMNRLTDDPRWLPAMSERVTRMVQDRN
HPSVIIWSLGNESGHGANHDALYRWIKSVDPSPRPVQYEGGGADTTATDIIICPMYARVD
EDQPFPAVPKWSIKKWLSPGETRPLILCEYAHAMGNSLGGFAKYWQAFRQYPRLQGG
FVWDVWDQSLIKYDENGPNWSAYGGDFGDTPNDRQFCMNGLVFADRTPHPALTEAKHQ
QOFFQFRLSGQTIEVTSEYLFRRHSDNELLHWMVALDGGKPLASGEVPLDVAPQGKQLIE
LPPLPQESAGQLWLTVRVVQPNATAWSEAGHISAWQQWRLAENLSVTLPAASHAIPH
LTTSEMDFCIELGNKRWFNRQSGFLSQMWIGDKKQLLTPLRDQFTRAPLDNDIGVSE
ATRIDPNAWVERWKAAGHYQAEALLQCTADTLADAVLITTAHAWQHOGKTLFISRKT
YRIDGSGQMAITVDVEVASDTPHPARIGLNCQLAQVAERVNWLGLGPQENYPDLRTAA
CFDRWDLPLSDMYTPYVFPSENGLRCTRELNYGPHQWRGDFQFNISRYSQQLMETS
HRHLLHAEEGTWNIDGFHMGIGGDDSWSPSVSAEFQLSAGRYHYQLVWCQK"
gene 4410..5663
/gene="lacY"
CDS 4410..5663
/gene="lacY"
/note="lactose permease"
/codon_start=1
/transl_table=11
/protein_id="AAA24054.1"
/db_xref="GI:146578"
/translation="MYYLKNTNFWMFGLFFFFFYFFIMGAYFPFFPIWLHDINHISKSD
TGIIFAAISLFSLLFQPLFGLLSDKLGLRKYLLWIIITGMLVMFAPFFIFIFGPPLLQYN
ILVGSIVGGIYLGFCFNAGAPAVEAFIEKVSRRSNFEFGRARMFGCVGWALCASIVGI
MFTINNQFVFWLGGSCALILAVLLFFAKTDAPSSATVANAVGANHSASFSLKLALELFR
QPKLWFLSLYVIGVSCTYDVFDQQFANFFTSFFATGEQGTRVFGYVTTMGELLNASIM
FFAPLIINRIGGKNALLLAGTMSVRIIGSSFATSALEVVLKTLHMFVFPFLLVGCF
KYITSQFEVRFSATIYLVCFCCFFKQLAMIFMSVLAGNMYESIGFQAYLVGLVALGF
TLISVFTLSGPGPLSLLRRQVNEVA"
gene 5727..6338
/gene="lacA"
CDS 5727..6338
/gene="lacA"
/note="thiogalactoside acetyltransferase (ttg start
codon)"
/codon_start=1
/transl_table=11
/protein_id="AAA24055.1"
/db_xref="GI:551814"
/translation="MNMPMTERIRAGKLFTDMCEGLPEKRLRGKTLMYEFNHSHPSEV
EKRESLIKEMFATVGENAWVEPPVYFSYGSNIHIGRNFYANFNLTIVDDYTVTIGDNV
LIAPNVTLSTGHVPVHHELKNGEMYSFPITIGNNVWIGSHVVINPGVTIGDNSVIGA
GSIVTKDIPPNVVAAGVPCRVIREDNRDKHYFFKDYKVESSV"
ORIGIN HindII site [Nature 274, 762-765 (1978)].
1 gacaccatcg aatggcgcaa aacctttcgc ggtatggcat gatagcgccc ggaagagagt

```
61 caattcagggt tggatgaatgt gaaaccagta acgttatatcg atgtcgcaga gtatgccggt
121 gtctcttatac agaccgtttc ccgcgtggtg aaccaggcca gccacgtttc tgcgaaaaacg
181 cgggaaaaaag tgggaagcggc gatggcgagg ctgaattaca tccccaacccg cgtggcacaa
241 caactggcgg gcaaacagtc gttgctgatt ggcgttgcca cctccagtct ggccctgcac
301 gcgcgctcgc aaattgtcgc ggcgattaaa tctcgcgcgc atcaactggg tgcagcgtg
361 gtggtgtcga tggtagaacg aagcggcgtc gaagcctgta aagcggcggg gcacaatctt
421 ctgcgcgaac gcgtcagtg gctgatcatt aactatccgc tggatgacca ggatgccatt
481 gctgtggaag ctgcctgcac taatgttccg gcgttatctt ttgatgtctc tgaccagaca
541 cccatcaaca gtattatctt ctcccatgaa gacggtagcg gactggcggt ggagcatctg
601 gtgcgattgg gtcaccagca aatcgcgctg ttagcggggc cattaagtctc tgtctcggcg
661 cgtctgcgtc tggctggctg gcataaatat ctactcgcga atcaaattca gccgatagcg
721 gaacgggaag gcgactggag tgccatgtcc ggttttcaac aaaccatgca aatgctgaat
781 gagggcatcg tccccactgc gatgctgggt gccaacgatc agatggcgct gggcgcaatg
841 cgcgccatta ccgagtcagg gctgcgcgtt ggtgcggata tctcggtagt gggatacgac
901 gataccgaag acagctcatg ttatatcccg ccgtcaacca ccatcaaaca ggattttcgc
961 ctgctggggc aaaccagcgt ggaccgcttg ctgcaactct ctcagggccca ggcggtgaag
1021 ggcaatcagc tgttgcccggt ctactgggtg aaaagaaaaa ccaccctggc gcccaatac
1081 caaaccgcct ctccccgcgc gttggccgat tcattaatgc agctggcacg acaggtttcc
1141 cgactggaaa gcgggcagtg agcgcaacgc aattaatgtg agttagctca ctcatagga
1201 accccagggt ttacacttta tgcttccggc tcgtatgttg tgtggaattg tgagcggata
1261 acaatttcac acaggaataa gctatgacca tgattacgga ttcactggcc gtcgttttac
1321 aacgtcgtga ctgggaaaac cctggcgcta cccaacttaa tcgccttgca gcacatcccc
1381 ctttcgccag ctggcgtaat agcgaagagg cccgcaccga tcgccttcc caacagttgc
1441 gcagcctgaa tggcgaatgg cgctttgcct ggtttccggc accagaagcg gtgcggaaa
1501 gctggctgga gtgcgatctt cctgaggcgc atactgtcgt cgtccctca aactggcaga
1561 tgcacggtta cgatgcgccc atctacacca acgtaacctc tccattacg gtcaatccgc
1621 cgtttgttcc cacggagaat ccgacgggtt gttactcgtc cacatttaat gttgatgaa
1681 gctggctaca ggaaggccag acgcgaatta ttttgatgg cgttaactcg gcgtttcatc
1741 tgtggtgcaa cgggcgctgg gtcggttacg gccaggacag tcgtttgccg tctgaatttg
1801 acctgagcgc atttttacgc gccggagaaa accgcctcgc ggtgatgggt ctgcgttga
1861 gtgacggcag ttatctggaa gatcaggata tgtggcggat gagcggcatt tccgtgacg
1921 tctcgttget gcataaaccc actacacaaa tcagcgattt ccatgttgcc actcgttta
1981 atgatgattt cagccgcgct gtactggagg ctgaagttca gatgtgcggc gagttgcgtg
2041 actacctacg ggtaacagtt tctttatggc aggtgaaac gcaggtcgcc agcggcaccg
2101 cgcctttcgg cggtgaaatt atcgatgagc gtggtggtta tgccgatcgc gtcacactac
2161 gtctgaacgt cgaaaacccg aaactgtgga gcgcgaaat cccgaatctc tatcgtgcg
2221 tgggtgaact gcacaccgcc gacggcacgc tgattgaagc agaagcctgc gatgtcgggt
2281 tccgcgaagt gcggtatgaa aatggtctgc tgcgtctgaa cggcaagccg ttgctgattc
2341 gaggcgttaa ccgtacagag catcatcctc tgcatggtca ggtcatggat gagcagcga
2401 tgggtgcagga tatcctgctg atgaagcaga acaactttaa cgccgtgcgc tgttcgcatt
2461 atccgaacca tccgtgtggt tacacgctgt gcgaccgcta cggcctgtat gtggtggatg
2521 aagccaatat tgaaaccac ggcatggtgc caatgaatcg tctgaccgat gatccgcgt
2581 ggctaccggc gatgagcgaa cgcgtaacgc gaatggtgca gcgcgatcgt aatcacccga
2641 gtgtgatcat ctggtcgtg gggaatgaat caggccacgg cgctaatac gagcgcgtgt
2701 atcgttggtt caaatctgtc gatccttccc gcccggtgca gtatgaaggc ggcggagccg
2761 acaccacggc caccgatatt atttgcccga tgtacgcgcg cgtggatgaa gaccagccct
2821 tcccggtgt gccgaaatgg tccatcaaaa aatggctttc gctacctgga gagacgcgcc
2881 cgctgatcct ttgcgaatac gccacgcga tgggtaacag tcttgccggt ttcgctaaat
2941 actggcaggc gtttcgtcag tatccccgtt tacaggggcg cttcgtctgg gactgggtg
3001 atcagtcgct gattaaatat gatgaaaacg gcaaccctg gtccggttac ggcggtgatt
3061 ttggcgatac gccgaacgat ccgagttct gtatgaacgg tctggtcttt gccgaccgca
3121 cgcgcgcatc agcgtgacg gaagcaaaac accagcagca gttttccag ttcggttat
3181 ccgggcaaac catcgaagt accagcgaat acctgttccg tcatagcgat aacgagctcc
3241 tgcactggat ggtggcgctg gatggtgaag cgctggcaag cggatgaagt cctctggatg
3301 tcgctccaca aggtaaacag ttgattgaac tgctgaact accgcagccg gagagcggc
3361 ggcaactctg gctcacgta cgcgtagtgc aaccgaacgc gaccgcattg tcagaagccg
3421 ggcacatcag gcctggcgag cagtggcgtc ttgcygaaaa cctcagtgtg acgctccccg
3481 ccgcgtccca cgccatcccg catctgacca ccagcgaaat ggatttttgc atcgagctgg
3541 gtaataagcg ttggcaattt aaccgccagt caggctttct ttcacagatg tggattggcg
3601 ataaaaaaca actgctgacg ccgctgcgcg atcagttcac ccgtgcaccg ctggataacg
3661 acattggcgt aagtgaagcg accgcattg accctaacgc ctgggtcgaa cgctggaagg
3721 cggcgggcca ttaccaggcc gaagcagcgt tgttgacgtg cacggcagat acacttgctg
3781 atgcggtgct gattacgacc gtcacgcgt ggcagcatca ggggaaaacc ttatttatca
3841 gccggaaaac ctaccgatt gatggtagt gtcaaatggc gattaccgtt gatgttgaag
```



```
3901 tggcgagcga tacaccgcat ccggcgcgga ttggcctgaa ctgccagctg gcgcaggtag
3961 cagagcgggt aaactggctc ggattagggc cgcaagaaaa ctatcccagc cgccttactg
4021 ccgcctgttt tgaccgctgg gatctgccaat tgtcagacat gtataccccg tacgtcttcc
4081 cgagcgaaaa cggctctgcg tgcgggacgc gcgaattgaa ttatggccca caccagtggc
4141 gcggcgactt ccagttcaac atcagccgct acagtcaaca gcaactgatg gaaaccagcc
4201 atcgccatct gctgcacgcg gaagaaggca catggctgaa tatcgacggt ttccatatgg
4261 ggattggtgg cgacgactcc tggagcccg tgcgtatcggc ggaattccag ctgagcgccg
4321 gtcgctacca ttaccagttg gtctggtgtc aaaaataata ataaccgggc aggccatgtc
4381 tgcccgatatt tcgctgaagg aaatccatta tgtactattt aaaaaacaca aacttttggg
4441 tgttcggttt attctttttc ttttactttt ttatcatggg agcctacttc ccgtttttcc
4501 cgatttggtc acatgacatc aaccatatca gcaaaagtga tacgggtatt atttttgccg
4561 ctattttctc gttctcgcta ttattccaac cgctgtttgg tctgctttct gacaaactcg
4621 ggctgcgcaa atacctgctg tggattatta ccggcatgtt agtgatgttt gcgccgttct
4681 ttatttttat cttcgggcca ctgttacaat acaacatttt agtaggatcg attgttggtg
4741 gtattttatc aggtttttgt tttaacgccg gtgcgccagc agtagaggca tttattgaga
4801 aagtcagccg tcgcagtaat ttogaatttg gtcgcgcgcg gatgtttggc tgtgttggtc
4861 gggcgctgtg tgcctcgatt gtcggcatca tgttcaccat caataatcag tttgttttct
4921 ggctgggctc tggctgtgca ctcatcctcg ccgttttact ctttttcgcc aaaacggatg
4981 cgccctcttc tgccacggtt gccaatgcgg taggtgccaa ccattcggca tttagcctta
5041 agctggcact ggaactgttc agacagccaa aactgtggtt tttgtcactg tatgttattg
5101 gcgtttcctg cacctacgat gtttttgacc aacagtttgc taatttcttt acttcgttct
5161 ttgctaccgg tgaacagggt acgcgggtat ttggctacgt aacgacaatg ggcgaattac
5221 ttaacgcctc gattatgttc tttgcgccac tgatcattaa tcgcatcggg gggaaaaacg
5281 cctcgctgct ggtcggcact attatgtctg tacgtattat tggctcatcg ttcgccacct
5341 cagcgctgga agtgggttatt ctgaaaacgc tgcataatgt tgaagtaccg ttcttgctgg
5401 tgggtgctt taaatatatt accagccagt ttgaagtgcg tttttcagcg acgatttattc
5461 tgggtctgtt ctgcttcttt aagcaactgg cgatgatttt tatgtctgta ctggcgggca
5521 atatgtatga aagcatcggg ttccaggggc cttatctggt gctgggtctg gtggcgctgg
5581 gcttcacctt aatttcctg ttacgcctta gcggccccgg cccgctttcc ctgctgcgtc
5641 gtcaggtgaa tgaagtcgct taagcaatca atgtcggatg cggcgcgacg cttatccgac
5701 caacatatca taacggagtg atcgcattga acatgccaat gaccgaaaga ataagagcag
5761 gcaagctatt taccgatatg tgcgaaggct taccggaaaa aagacttcgt gggaaaacgt
5821 taatgtatga gtttaatac tcgcatccat cagaagttga aaaaagagaa agcctgatta
5881 aagaaatgtt tgccacggtg ggggaaaaacg cctgggtaga accgcctgtc tatttctctt
5941 acggttccaa catccatata ggccgcaatt tttatgcaaa tttcaattta accattgtcg
6001 atgactacac ggtacaacatc ggtgataacg tactgattgc acccaacgtt actctttccg
6061 ttacgggaca ccctgtacac catgaattga gaaaaaacgg cgagatgtac tcttttccga
6121 taacgattgg caataacgtc tggatcggaa gtcattgtgt tattaatcca ggcgtcacca
6181 tcggggataa ttctgttatt ggcgcggtta gtatcgtcac aaaagacatt ccaccaaacg
6241 tctggtgggc tggcgcttct tgtcgggtta ttgcgaaat aaacgaccgg gataagcact
6301 attatttcaa agattataaa gttgaatcgt cagtttaaat tataaaaatt gcctgatacg
6361 ctgcgcttat caggcctaca agttcagcga tctacattag ccgcatccgg catgaacaaa
6421 gcgcaggaac aagcgtcgca tcatgcctct ttgaccaca gctgcggaaa acgtactggt
6481 gcaaaacgca gggttatgat catcagccca acgacgcaca gcgcatgaaa tgcccagtcc
6541 atcaggtaat tgccgctgat actacgcagc acgccagaaa accacggggc aagcccggcg
6601 atgataaaac cgattccctg cataaacgcc accagcttgc cagcaatagc cggttgaca
6661 gagtgatcga gcgccagcag caaacagagc ggaaacgcgc cgcccagacc taaccacac
6721 accatcgccc acaataccgg caattgcatc ggcagccaga taaagccgca gaacccacc
6781 agttgtaaca ccagcgccag cattaacagt ttgcgccgat cctgatggcg agccatagca
6841 ggcacagca aagctcctgc ggcttgccca agcgtcatca atgccagtaa ggaaccgctg
6901 tactgcgcgc tggcaccaat ctcaatatag aaagcgggta accaggcaat caggctggcg
6961 taaccgccgt taatcagacc gaagtaaaca ccagcgtcc acgcgcgggg agtgaatacc
7021 acgcgaaccg gagtgttgt tgtcttgtgg gaagaggcga cctcgcgggc gctttgccac
7081 caccaggcaa agagcgcaac aacggcaggc agcgcacca ggcgagtgtt tgataaccag
7141 tttcgctatg ttgaactaac caggcggtta tggcggcacc aagcccaccg ccgccatca
7201 gagccgcgga ccacagcccc atcaccagtg gcgtgcgctg ctgaaaccgc cgtttaatca
7261 ccgaagcatc accgcctgaa tgatgccgat cccacccca ccaagcagtg cgctgctaag
7321 cagcagcgca ctttgcgggt aaagctcacg catcaatgca ccgacggcaa tcagcaacag
7381 actgatggcg acactgcgac gttcgtcgac atgctgatga agccagcttc cggccagcgc
7441 cagcccgcgc atggttaacca ccggcagagc ggtcgac
```

//

Jan 29 2004 07:45:51